MOD $p$ DESCENT FOR HILBERT MODULAR FORMS

CHANDRASHEKAR KHARE

1. Introduction

Let $F$ be a totally real extension of $\mathbb{Q}$ and

$$\rho_F : \text{Gal}(\overline{F}/F) \rightarrow \text{GL}_2(k)$$

be an absolutely irreducible, continuous, and odd representation, with $k$ a finite field of characteristic $p > 2$, and where $G_F := \text{Gal}(\overline{F}/F)$ is the absolute Galois group of $F$. For a set of places $S$ of $F$, we shall denote by $G_{F,S}$ the Galois group of the maximal extension of $F$ in $\overline{F}$ that is unramified outside $S$.

By an odd representation we mean $\det(\rho_F(c)) = -1$, where $c$ runs through the conjugacy classes of complex conjugation with respect to each real place of $F$. In this situation, the natural extension of Serre’s conjecture (cf. [S]) would say that $\rho_F$ is modular, i.e., arises from a holomorphic Hilbert modular form of some weight and level. As usual by arises from we mean that it is the reduction mod $\wp$ of the $\wp$-adic Galois representation associated to a holomorphic Hilbert modular form of weight $(\ell, \cdots, \ell)$ for an integer $\ell \geq 2$, with $\wp$ a place of $\mathbb{Q}$ above $p$.

Now assume that $\rho_F$ is modular and extends to a representation $\rho_Q$ of the absolute Galois group $G_Q$ of $\mathbb{Q}$, i.e.,

$$\rho_Q \rightarrow \text{GL}_2(k'),$$

where $k'$ is a finite extension of $k$ and $\rho_Q|_{G_F}$ is isomorphic to $\rho_F$ in $\text{GL}_2(k)$ or $\text{GL}_2(k')$: the choice of $k$ or $k'$ does not matter as, for instance, the Brauer group of a finite field is trivial. This extension need not be unique, but any such extension is odd and absolutely irreducible.

Under these circumstances the following question is natural and was raised in [K]:

**Question 1. Is the extension $\rho_Q$ modular?**

Assuming Serre’s conjecture (cf. [S]) the answer to this question is in the affirmative as $\rho_Q$ is odd and absolutely irreducible.

It is the purpose of this paper to provide an answer to this question, under some particular circumstances, using the work of [F] and the trilogy [SW], [SW1], [SW2] on the one hand, and [L] and [R] on the other. The simple argument used here is reminiscent of [C] and [HM], i.e., we use the result of [R] on lifting Galois

Received May 15, 2000.
representations, and then go via results of [F] and [SW], [SW1], [SW2] to the automorphic or Hecke side. Namely we use:

- The identification of Hecke and deformation rings proved in the preprints [F] and [SW], [SW1], [SW2] for totally real fields in many cases. As none of these preprints is available to us, we will simply state a result (Theorem 3) that we shall assume to be true, and which should be proven in either of [F] or [SW], [SW1], [SW2].
- The lifting result of [R].
- The characterisation of the image of base change for cyclic, prime degree extensions in [L].

2. The assumptions and the theorem

The particular circumstances are the following:

1. Assume that $\rho_F$ is (absolutely) irreducible, odd and modular, i.e., $\rho_F$ arises from a Hilbert modular form.
2. Assume that $F/\mathbb{Q}$ is a solvable extension, $p$ does not divide the class number of $F$, and is unramified in $F$.
3. Assume that $\rho_F$ restricted to $F(\sqrt{p^*})$ is absolutely irreducible where $p^* = (-1)^{p-1/2}$.p.
4. Assume that the order of the projective image of $\rho_Q$ is divisible by $p$, that the restriction $\rho_Q$ to $D_p$, the decomposition group at $p$, is of the form

$$\begin{pmatrix} \varepsilon & * \\ 0 & \delta \end{pmatrix},$$

with $\varepsilon$ and $\delta$ ramified and unramified characters of $D_p$, respectively.

The modularity of $\rho_F$ and the solvability of $F/\mathbb{Q}$ are essential conditions for our method to work, while the other conditions are of a more technical nature and required to apply the work of [F] and [SW], [SW1] and [SW2].

The following theorem is the main result of this paper:

**Theorem 1.** Assuming the conditions above, the extended representation $\rho_Q$ is modular.

2.1. Proof of theorem. The following theorem of [R] is crucial for us.

**Theorem 2.** Suppose $\rho_Q : G_{Q,S} \to \text{GL}_2(k')$ is odd and $\text{Ad}^0(\rho_Q)$ is absolutely irreducible. Suppose further that $\rho_Q|_{D_p}$ is ordinary and ramified, and the order of the projective image of $\rho_Q$ is divisible by $p$. Then there is a finite set of primes $R$ containing $S$ and a lift $\tilde{\rho}_Q : G_{Q,R} \to \text{GL}_2(W(k'))$ of $\rho_Q$ with $\tilde{\rho}_Q|_{D_p}$ of the form

$$\begin{pmatrix} \chi^{\ell-1}\varepsilon' & * \\ 0 & \delta' \end{pmatrix},$$

with $\varepsilon'$ a finite order character and $\delta'$ an unramified character of $D_p$, and the determinant of $\tilde{\rho}_Q = \chi^{\ell-1}\psi$ for $\psi$ some finite order character, and $\chi$ the $p$-adic cyclotomic character and $\ell$ an integer $\geq 2$. 
Remark. Note that the lift $\tilde{\rho}_Q$ need not be “minimally ramified”. Note that the results of [R] are stated for $p > 5$, but the methods of [R] can be modified to work for $p > 2$: for $p = 5$ see [T], and for $p = 3$ see Section 6.2 of [KR].

The reducibility of $Ad^0(\rho_Q)$, together with the irreducibility of $\rho_Q$, implies that $\rho_Q$ is dihedral. If $\rho_Q$ is dihedral and odd, the modularity of $\rho_Q$, and hence Theorem 1 follows from a classical result of Hecke. Thus we can and will assume that we are in the case when $Ad^0(\rho_Q)$ is absolutely irreducible.

We will consider isomorphism classes of representations such as $\tilde{\rho}_Q$ above, when considered as representations into $GL_2(W(k') \otimes \mathbb{Q}_p)$ rather than into $GL_2(W(k'))$. Though we are interested in the reduction mod $p$ of such representations, as we are assuming that the mod $p$ representations considered are irreducible, after the theorem of Brauer-Nesbitt that reduction mod $p$ is well-defined up to semisimplification, this weaker notion of isomorphism is good enough for us.

Now we quote the (assumed) result of [F] and [SW], [SW1] and [SW2] needed here.

**Theorem 3.** The representation $\tilde{\rho}_Q|_{G_F}$ is associated to a unique Hilbert modular newform $f$ over $F$ of weight $(\ell, \cdots, \ell)$, $\ell \geq 2$.

**Remark.** By the word “associated” we mean that $\tilde{\rho}_Q|_{G_F}$ is the $p$-adic representation attached to $f$ (for some place $\wp$ above $p$) by results of Taylor, Blasius-Rogawski et al.

“Proof”. As we said in the introduction we merely assume this. To say a few words about a possible proof nevertheless, using [F] and [SW], [SW1], [SW2] (see Theorem 7.1 of [E] and the discussion after it for relevant information), this presumably follows from

- Assumptions 1 to 4, which in particular imply that $\rho_F|_{D_{\wp}}$ for any place $\wp$ of $F$ above $p$ is ordinary with distinct characters on the diagonal
- $\tilde{\rho}_Q|_{G_F}$ is ordinary at any prime above $p$.

### 2.2. Proof of Theorem 1.

As $\tilde{\rho}_Q|_{G_F}$ is invariant under the conjugation action of $\text{Gal}(F/\mathbb{Q})$, we deduce that $f^\sigma = f$ for all $\sigma \in \text{Gal}(F/\mathbb{Q})$.

Let

$$1 = G^{(i)} \leq G^{(i-1)} \cdots \leq G^{(0)} = \text{Gal}(F/\mathbb{Q})$$

be a composition series for $\text{Gal}(F/\mathbb{Q})$ with $G^{(j)}$ the commutator subgroup of $G^{(j-1)}$ for $1 \leq j \leq i$: thus the successive quotients $G^{(j)}/G^{(j+1)}$ are abelian. Denote the subfields of $F$ that are the fixed fields of $G^{(j)}$ by $F_{i-j} = F_0$ for $j = 0, \cdots, i$; thus $F_i = \mathbb{Q}$ and $F_0 = F$. Note that these subfields are Galois extensions of $\mathbb{Q}$.

Let $F_{0,k} = F$ for $0 \leq k \leq r$ be the subextensions of the abelian extension $F_0/F_1$, such that $F_{0,0} = F$, $F_{0,r} = F_1$, and $F_{0,s}/F_{0,s+1}$ is a cyclic extension of prime degree for $0 \leq s \leq r-1$. By the results of [L], as $f_0 := f$ is invariant under the action of $\text{Gal}(F_{0,0}/F_{0,1})$, it arises by “base change” with respect to the prime, cyclic extension $F_{0,0}/F_{0,1}$ from a Hilbert modular form $f_{0,1}$ of weight $(\ell, \cdots, \ell)$ ($\ell \geq 2$) for $F_{0,1}$. 

This means that the Galois representation \( \rho_1 \) of \( G_{F_{0,1}} \) associated to \( f_{0,1} \) when restricted to \( G_{F_{0,0}} \) is isomorphic to the restriction of \( \tilde{\rho}_Q \) to \( G_{F_{0,0}} \). But as we have assumed that \( \rho_{F_{0,0}} \) (i.e., \( \rho_F \)) is absolutely irreducible, and \( F_{0,0}/F_{0,1} \) is cyclic, it follows that any representation of \( G_{F_{0,1}} \) which restricts to \( G_{F_{0,0}} \) to give \( \tilde{\rho}_Q|_{G_{F_{0,0}}} \) is isomorphic to a twist, by a character of Gal \((F_{0,0}/F_{0,1}), \) of \( \tilde{\rho}_Q \) restricted to \( G_{F_{0,1}} \). Thus we deduce that \( \tilde{\rho}_Q|_{G_{F_{0,1}}} \) is itself modular. Noting that \( \tilde{\rho}_Q|_{G_{F_{0,1}}} \) is invariant under the action of Gal \((F_{0,1}/F_{0,2}), \) we argue as before, using the results of [L] applied to the cyclic prime degree extension \( F_{0,1}/F_{0,2}, \) to show that \( \tilde{\rho}_Q|_{G_{F_{0,2}}} \) is modular. Iterating the argument we conclude that \( \tilde{\rho}_Q|_{G_F} \) is modular.

Continuing in this way we deduce successively that \( \tilde{\rho}_Q|_{G_{F_j}} \) is modular for \( j = 0, \ldots, i-1 \) and hence finally that \( \tilde{\rho}_Q \) is modular of some weight \( \ell \geq 2. \) Thus \( \rho_Q, \) the mod \( p \) reduction of \( \tilde{\rho}_Q, \) is itself modular and this finishes the proof of Theorem 1.

**Remark.** Note that the results of [L] are for prime degree, cyclic extensions, while above we have availed of Galois representations to argue in the solvable case.

### 3. Descent for cyclic extensions

In the case \( F/Q \) is a totally real, cyclic extension of prime degree we can in many cases start with descent data on \( \rho_F, \) rather than conditions on \( \rho_Q, \) just as in the classical case (cf. [L]). We note the following consequence of the proof above.

**Theorem 4.** Let \( F/Q \) be a totally real, cyclic extension of prime degree, coprime to the cardinality of \( k^*, \) with \( k \) a finite field of characteristic \( p, \) \( p \) unramified in \( F, \) and \( \rho_F : G_F \to GL_2(k) \) an irreducible, modular, mod \( p \) representation that is invariant under the action of Gal \((F/Q), \) such that the order of its projective image is divisible by \( p, \) with determinant the mod \( p \) cyclotomic character, and satisfying the assumptions 1 to 3 above.

Then it extends to a representation \( \rho_Q : G_Q \to GL_2(k') \) for some finite extension \( k' \) of \( k. \) Further if we suppose that \( \rho_F \) restricted to the decomposition group at a place \( \varphi \) of \( F \) above \( p \) is ordinary, i.e., has an unramified quotient, and that \( F \) has class number prime to \( p, \) then the extended representation \( \rho_Q \) is modular.

**Proof.** That \( \rho_F \) as in the theorem extends to a representation of \( G_Q \) follows from standard cohomological arguments. Namely, the obstruction lies in \( H^2(G_Q, k^*), \) and it vanishes on restriction to \( H^2(G_F, k^*). \) But as \( F/Q \) is coprime to the order of \( k^*, \) we deduce that the the obstruction in \( H^2(G_Q, k^*) \) itself vanishes, and hence that the representation \( \rho_F \) extends to \( G_Q. \) We pick an extension, say \( \rho_Q. \) It is unique up to twisting by a character of Gal \((F/Q). \)

Under the further hypothesis that the restriction of \( \rho_F \) to the decomposition group of a place of \( F \) above \( p \) is ordinary, we claim that \( \rho_Q|_{D_p} \) is again ordinary. This follows as, if the semisimplification of \( \rho_F \) restricted to the decomposition group at a place of \( F \) above \( p \) is the sum of two characters \( \psi \) and \( \psi', \) with say
ψ' unramified, then the determinant of ρ_F being ramified at all places above p implies that both the characters ψ and ψ' extend to $G_\mathbb{Q}$. As p is unramified in F, any extension of ψ' to $G_\mathbb{Q}$ is unramified. This justifies the claim.

The assumption that $F/\mathbb{Q}$ is unramified at p, together with that on the determinant character of ρ_F, implies that ρ_Q is not in the excluded case of Theorem 1 part (b) of [R]. After this we are done by the same argument that was applied when considering $F_{0,0}/F_{0,1}$ in the proof of Theorem 1.

4. Rationality and compatibility of lifts

Let $\rho : G_\mathbb{Q} \to \text{GL}_2(k)$ be an odd, irreducible representation with k a finite field of characteristic $p > 3$. Fix an auxiliary odd and surjective representation $\rho_3 : G_\mathbb{Q} \to \text{GL}_2(\mathbb{Z}/3\mathbb{Z})$ that is ordinary at 3, irreducible on restriction to $\mathbb{Q}(\sqrt{-3})$. By Wiles, any ordinary lifting of $\rho_3$, such that its determinant character is "arithemtic", i.e., a positive integral power of the 3-adic cyclotomic character up to a finite order character, is modular: to see this note that, by the Langlands-Tunnell theorem, $\rho_3$ is modular. Thus if we could find a lifting of $\rho$ that is "compatible" with an ordinary 3-adic lifting of $\rho_3$ with "arithemtic" determinant character, we would have proven Serre’s conjecture: note that after Wiles the latter lifting is algebraic (i.e., the traces of frobenii are in a number field) and hence it makes sense to talk of compatibility. The results of [R] do not a priori give algebraic liftings: but see a very recent preprint of Taylor (cf., [T1]) that proves algebraicity of the liftings produced in [R] in a large number of cases.

Notice that we have a lot of freedom in the choice of $\rho_3$ and thus could choose it so that the corresponding deformation rings are unobstructed, though it is not clear if this will be helpful. On the other hand, the choice of $\rho_3$ cannot be unconstrained if we hope to succeed, as $\rho_3$ should arise from newforms of level divisible by the conductor of $\rho$. These considerations give rise to the following 2 natural questions:

Question 2. Let p and q be distinct primes, $\rho_p : G_\mathbb{Q} \to \text{GL}_2(\mathbb{F}_p)$, $\rho_q : G_\mathbb{Q} \to \text{GL}_2(\mathbb{F}_q)$ be irreducible, continuous representations. Suppose $\rho_p$ (resp., $\rho_q$) is semistable at all primes different from p (resp., q) and finite, flat at p (resp., q), and $\det(\rho_p)$ (resp., $\det(\rho_q)$) is the mod p (resp., mod q) cyclotomic character. Suppose further that $\rho_p$ and $\rho_q$ have the same prime to residue characteristic Artin conductor. Then do there exist compatible lifts of $\rho_p$ and $\rho_q$?

Question 3. Let p and q be distinct primes, $\rho_p : G_\mathbb{Q} \to \text{GL}_2(\mathbb{F}_p)$, $\rho_q : G_\mathbb{Q} \to \text{GL}_2(\mathbb{F}_q)$ irreducible representations that satisfy Serre’s conjecture, i.e., which are irreducible and arise from modular forms. Suppose $\rho_p$ (resp., $\rho_q$) is semistable at all primes different from p (resp., q) and finite, flat at p (resp., q), and $\det(\rho_p)$ (resp., $\det(\rho_q)$) is the mod p (resp., mod q) cyclotomic character. Suppose further that $\rho_p$ and $\rho_q$ have the same prime to residue characteristic Artin conductor. Then does there exist a newform of some level which gives rise to $\rho_p$ and $\rho_q$ simultaneously?
The local conditions are just to ensure that there are no local constraints in answering the question affirmatively, and may be refined further. Note that the conditions on $\rho_p$ (resp., $\rho_q$) imply that it’s determinant is the mod $p$ (resp., mod $q$) cyclotomic character. By mod $p$ and mod $q$ representations above we mean those that arise after having fixed a prime above $p$ and $q$.

One of the problems with Question 3 is that it is difficult to predict levels at which to find the desired form that gives rise to $\rho_p$ and $\rho_q$: for example not just any level from which $\rho_p$ and $\rho_q$ arise individually will work. To see this, we claim that if $S_2(\Gamma_0(N))$ has dimension greater than 1, for an integer $N \geq 1$, then for large enough pairs of primes $p$ and $q$, representations $\rho_p$ and $\rho_q$ which arise individually from $S_2(\Gamma_0(N))$, do not arise compatibly from $S_2(\Gamma_0(N))$.

To justify the claim, if $f_1, f_2, \cdots, f_i \in S_2(\Gamma_0(N))$ are the distinct newforms, then pick large enough primes $p$ and $q$ so that the mod $p$ (resp., mod $q$) representations arising from the $f_i$’s are mutually distinct and irreducible, and have conductor $N$. Pick $\rho_p$ to be the mod $p$ representation arising from $f_1$, and $\rho_q$ to be the mod $q$ representation arising from $f_2$. Then $\rho_p$ and $\rho_q$ do not arise simultaneously from any of the $f_i$’s, justifying the claim. B. Mazur and K. Ribet have considered a question similar to Question 3 above, and there are some computations of William Stein related to it (see [St]).

For 1-dimensional characters the analogous questions have affirmative answers if and only if there are no local constraints. Namely, we claim that if we have characters $\eta_p : G_{\mathbb{Q}} \rightarrow \mathbb{F}_{p}^\ast$ (resp., $\eta_q : G_{\mathbb{Q}} \rightarrow \mathbb{F}_{q}^\ast$) then there exists a character $\eta : G_{\mathbb{Q}} \rightarrow \mu_\infty$, where $\mu_\infty$ is the torsion subgroup of $\mathbb{Q}^\ast$, which reduces modulo a chosen place above $p$ (resp., above $q$) to $\eta_p$ (resp., $\eta_q$) if and only if this can be done locally at all primes. This follows because a necessary and sufficient condition for the existence of $\eta$ is that if $\tilde{\eta}_p$ (resp., $\tilde{\eta}_q$) is a lifting of $\eta_p$ (resp., $\eta_q$) with the same order (a Teichmuller lifting) then an $\eta$ of the desired sort exists if and only if the order of $\tilde{\eta}_p \tilde{\eta}_q^{-1}$ is of the form $p^\alpha q^\beta$. This together with the Cebotarev density theorem justifies the claim.

5. Some remarks

- There are many instances in which one knows ordinary at $p$ liftings of $\rho_{\mathbb{Q}}$ exist simply because the corresponding ordinary (at $p$) deformation ring (with respect to some ramification data) is unobstructed (see [M] and [BM] for examples of unobstructed ordinary deformation rings). In these situations we evidently do not need the result of [R].

- Some examples of non base change forms for totally real fields $F$ such that they satisfy descent data mod $p$ can be constructed by twisting base change forms by certain characters congruent to the identity mod $p$. One can give a class of less trivial examples for our main theorem as follows:

  Consider an odd representation $\rho_{\mathbb{Q}} : G_{\mathbb{Q}} \rightarrow GL_2(\mathbb{F}_p)$ that has determinant the mod $p$ cyclotomic character, is surjective, with $p$ congruent to 1 mod 4, with $p \geq 5$, and $F$ a real quadratic field. By the Cebotarev density
theorem, there is a set of primes \( \{q\} \) of positive density such that \( \rho_Q|D_q \) is of the form

\[
\begin{pmatrix}
-xq & 0 \\
0 & x
\end{pmatrix}
\]

with \( x \in F_p \) such that \( x^2 \) congruent to -1 mod \( p \), and such that \( q \) is inert in \( F \). Thus the ratio of the eigenvalues of \( \rho_F(\text{Frob}_Q) \) is the image of \( q^2 = \text{card}(O_F/Q) \) in \( F_p \), where \( Q \) is the prime of \( F \) above \( q \). Assume that \( \rho_Q|G_F \) arises from a Hilbert modular form. Then Ribet’s raising levels arguments (cf. [Ri]), applied after switching to a definite quaternion algebra over \( F \) that is unramified at all finite places, yield that \( \rho_F \) arises from a form that is Steinberg at \( Q \). From this we see easily that such a form cannot be a base change form, as otherwise \( \rho_Q \) would itself arise from a form that is Steinberg at \( q \). This cannot happen because of our assumption that the ratio of the eigenvalues of \( \rho_Q(\text{Frob}_q) \) is \(-q\).

6. Conclusion and acknowledgements

- The author had observed in the summer of 1997 that results about \( p \)-adic semistable (at \( p \)) lifting of representations such as \( \rho_Q \) above, results of the type proved in [F], together with the classical base change theory of [L] would prove certain cases of mod \( p \) descent for Hilbert modular forms. Because of [R], this strategy can now be implemented, resulting in the present work.
- The method of proof of Thorem 1 has found an application to proving more cases of Artin’s conjecture in a recent preprint of R. Taylor (cf., [T]).
- The author thanks the National University of Singapore for its hospitality and financial support during the summer of 1999 when this paper was written. He also thanks Barry Mazur for helpful correspondence.

References


[St] W. Stein, *Serre’s conjecture mod pq?*, available at http://shimura.math.berkeley.edu/~was/Tables/serremodpq/


School of Mathematics, TIFR, Homi Bhabha Road, Mumbai 400 005, INDIA.

E-mail address: shekhar@math.tifr.res.in